

Genetic Distance among Selected Cotton Genotypes and Its Relationship with F₂ Performance

O. A. Gutiérrez, S. Basu, S. Saha,* J. N. Jenkins, D. B. Shoemaker, C. L. Cheatham, and J. C. McCarty, Jr.

ABSTRACT

Knowledge of genetic diversity and relationships among breeding materials has a significant impact on crop improvement. Association between parental divergence and progeny performance has not been well documented in cotton (*Gossypium hirsutum* L.). The objectives of this study were to estimate genetic diversity among selected cotton genotypes on the basis of simple sequence repeat (SSR) markers, and to investigate the relationship between genetic diversity and F₂-bulk population performance. Five U.S. and four Australian cultivars, and two day-neutral converted lines of *G. hirsutum* were genotyped by means of 90 SSR primer pairs providing 69 polymorphic marker loci. Genetic distance (GD) between genotypes ranged from 0.06 to 0.34 for the 11 parental genotypes. The highest GD (0.34) was observed between ST474 and the day-neutral converted line B1388. The lowest GD (0.06) was detected between cultivars FM832 and FM975. The GD between day-neutral converted lines and cultivars ranged from 0.26 to 0.34. Among the Australian cultivars, GD ranged from 0.06 to 0.19 while GD among U.S. cultivars varied from 0.10 to 0.22, indicating a narrow genetic base. Significant correlations between agronomic and fiber traits of F₂-bulk populations and GD ranged from negative to positive depending on the traits, genetic background, and environment. On the basis of SSR markers, GD revealed a lack of genetic diversity among all genotypes and it was a poor predictor of overall F₂ performance. However, when genotypes with maximum range of GD were present, it was a better predictor for some traits.

INFORMATION about the degree and distribution of genetic diversity and relationships among breeding materials has a significant effect on crop improvement. Selection of suitable parents is one of the most important criteria used to allocate resources to the most promising crosses and increase the efficiency of breeding programs. Molecular markers increasingly play an important role in crop improvement programs. They have been used to predict genetic variance among inbred lines (Manjarrez-Sandoval et al., 1997), estimate genetic diversity in crops (Wendel et al., 1992; Tatineni et al., 1996), protect plant cultivar rights (Smith and Smith, 1992), classify heterotic groups (Dudley et al., 1991; Senior et al., 1998), study phylogenetic relationships among crops and their wild relatives (Li et al., 2000), analyze pedigrees (Smith et al., 1997), and select desired traits (Young, 1999).

Simple sequence repeats, also called microsatellites,

are tandem repetitive DNA sequences. The availability and abundance of SSR markers throughout the cotton genome, their polymorphic nature, codominance, and polymerase chain reaction (PCR)-based assay make SSRs useful in detecting genetic diversity (Reddy et al., 2001).

The usefulness of GD as a predictor of hybrid performance has been studied in several crops. In maize (*Zea mays* L.) and sunflower (*Helianthus annuus* L.), significant correlations between GD and hybrid performance were observed by Lee et al. (1989), Smith et al. (1990), Lanza et al. (1997), and Cheres et al. (2000). In contrast, Godshalk et al. (1990) and Dudley et al. (1991) observed weak correlations between marker genotype and hybrid performance in maize. Furthermore, Martin et al. (1995) working with wheat (*Triticum aestivum* L.) found no association between measures of diversity and hybrid performance. Meredith and Brown (1998) using restriction fragment length polymorphic (RFLP) markers reported that in cotton the correlations between yield of F₂ hybrids, heterosis, and GD were very low ($r = 0.07$).

The utilization of heterosis for lint yield and fiber quality in F₂ hybrids has eluded most researchers because of inconsistent expression of heterosis and lack of economically effective means of delivering F₂ seeds to growers. Nonetheless, it has been demonstrated that F₂ hybrids can exhibit superior performance for agronomic and fiber traits when compared with their parental lines. Meredith (1990) reported that F₂ hybrids yielded 8% higher than their parents and no differences in adaptation were detected; however, the F₂ hybrids had significantly shorter fiber. Moreover, Tang et al. (1992, 1993a,b) observed that almost all F₂ hybrids developed from crosses between selected pest-resistant germplasms and cultivars were equal to their highest parent for most fiber traits. Perhaps the greatest motivation to investigate prediction of F₂ performance is to spur cotton improvement. Presently, F₂-bulk population performance is used in about 24% of commercial cotton breeding programs in the USA for early generation testing and identification of populations upon which to focus selection (Bowman, 2000).

The objectives of this study were to estimate GD on the basis of SSR markers among selected cotton genotypes, and to investigate the association between GD and fiber and agronomic characteristics of F₂-bulk populations.

MATERIALS AND METHODS

Plant Materials and Field Experiments

Eleven cotton genotypes, five U.S. cultivars, four cultivars developed by the CSIRO cotton breeding program in Australia,

O.A. Gutiérrez, S. Saha, J.N. Jenkins, and J.C. McCarty, Jr., USDA-ARS, Crop Science Res. Lab., Genetics and Precision Agriculture Research Unit, Mississippi State, MS 39762; S. Basu, Dep. of Plant and Soil Sciences, Mississippi State Univ., Mississippi State, MS 39762; C.L. Cheatham, USDA-ARS, Small Fruit Res. Stn., Poplarville, MS 39470; D.B. Shoemaker, Delta and Pine Land Company, Scott, MS 38772. Contribution of the USDA-ARS in cooperation with the Mississippi Agric. and Forestry Exp. Stn. Journal paper No. J 9891. Received 9 July 2001. *Corresponding author (ssaha@msa-msstate.ars.usda.gov).

Table 1. Pedigrees and origin of cotton cultivars and day-neutral converted lines used to estimate genetic distance and its relationship with F₂-bulk population performance.

Genotypes	Origin	Pedigree
A239	Guatemala	(DP16/PI 163693)//PI 163693*3
B1388	Nicaragua	DES56/PI 415112
DP50	Delta & Pine Land Co. (US)	DP16//DP Smoothleaf/DP45/3/DES56
DP90	Delta & Pine Land Co. (US)	DP6516/DP6582
FM832	Cotton Seed International (AU)	Confidential
FM963	Cotton Seed International (AU)	Confidential
FM975	Cotton Seed International (AU)	Confidential
FM989	Cotton Seed International (AU)	Confidential
PM1560	Paymaster Cottonseed (US)	DES119/La434-RKR
SG501	Sure Grow Seed, Inc. (US)	DES119/DES237-7
ST474	Stoneville Pedigreed Seed Co. (US)	DES119/ST453

lia, and two day-neutral converted lines of *G. hirsutum* derived from photoperiodic wild accessions were selected for this study (Calhoun et al., 1994) (Table 1). The U.S. cultivars are a representative sample of those grown in the Delta region and some represent the most influential breeding programs in terms of genetic contributions to modern cultivars (Bowman et al., 1996). The Australian cultivars were obtained from Cotton Seed International. These four were used because they were offered for sale to growers in the USA and their GD from U.S. cultivars was unknown.

Day-neutral converted line A239 was derived from photoperiodic primitive accession T239 (PI 163693) *G. hirsutum* race *latifolium* from Guatemala. A239 was developed from a cross of 'Deltapine 16' by T239 and selected for the day-neutral flowering habit in the F₂. A day-neutral selection was then backcrossed three times to T239 and selected for day neutrality after each backcross. Following the third backcross a single high fiber strength plant was selected in the BC₃F₂. Subsequently, a single high fiber strength plant was selected in the BC₃F₃ and BC₃F₄. The BC₃F₄ plant was advanced by bulk increase to the BC₃F₉. Day-neutral converted line B1388 was derived from accession T1388 (PI 415112) of unknown race from Nicaragua. Line B1388 was developed from a cross of 'DES 56' by T1388. In the F₂, a single high fiber strength day-neutral plant was selected. Subsequently, a single high fiber strength plant was selected in the F₃ and F₄. The F₄ plant was advanced by bulk increase to the F₉.

The mating design used for this experiment was an 11 parent half diallel. Crosses were made during 1998 at the R. R. Foil Plant Science Research Center, at Mississippi State University. The F₁ generation was grown in the USDA-ARS winter nursery at Tecoman, Colima, Mexico, to produce F₂ seed. Seed from the 54 F₂ hybrids (one cross was lost) and the eleven parents were grown at the Plant Science Research Center, Mississippi State, MS, in the summers of 1999 and 2000 at two locations which differ in soil type and rainfall pattern.

The experimental design was a randomized complete block design with four replications within each location. The experiments were planted on 11 and 12 May 1999, and 11 and 15 May 2000. Trials were planted in 9- or 12-m single-row plots with 10 plants m⁻¹ of row and a row spacing of 0.97 m. The soil type at the first location was a Leeper silty clay loam (fine, smectitic, nonacid, thermic, Vertic Epiaquepts) and at the second location was a Marietta loam (fine-loamy, siliceous, active, thermic, Fluvaquentic Eutrudepts) soil. Standard cultural practices were followed during the growing season.

Plots were harvested with a mechanical picker for yield determination on 7 and 8 Oct. 1999 and 20 Sept. and 10 Oct. 2000. Samples containing 50 bolls were hand-harvested both years from each plot before mechanical harvesting. The boll samples were weighed and ginned on a 10-saw laboratory gin for lint percentage calculations (100 × lint weight/seed cotton weight). A lint sample (20 g) from each boll sample repre-

sented a fiber sample. Fiber samples were sent to STARLAB Inc. (Knoxville, TN) for measurements of elongation (%), fiber strength (kN m kg⁻¹), 2.5% span length (mm), 50% span length (mm), and micronaire reading by single-instrument testing.

SSR Markers

Leaves were collected from field plots in summer 1999. Samples of 20 young leaves from each of the 11 parents were collected and freeze dried (Saha et al., 1997). The freeze-dried tissue was then ground into dry powder and stored at -20°C. Freeze-dried, powdered tissue was ground further with liquid nitrogen just before extraction following the manufacturer's protocol. DNA was isolated from 40 mg (dry weight) of cotton leaf tissue with the DNeasy Plant mini kit (Qiagen, Santa Clarita, CA) following the manufacturer's protocol. The DNA concentration was measured with a spectrophotometer. The purity of the uncut DNA samples was also visually determined by means of agarose gel electrophoresis.

The 90 SSR primer pairs were received from Research Genetics Inc., Huntsville, AL. The PCR was performed with 80 ng DNA as template, 0.15 μM each, fluorescent-labeled and nonlabeled specific SSR primer pairs labeled either HEX (4,7,2',4',5',7'-hexachloro-6-carboxyfluorescein), NED (7'8'-benzo-5'fluoro-2',4,7-trichloro-5-carboxyfluorescein), or FAM (6-carboxyfluorescein), 0.2 mM each dNTP, 1× GeneAmp PCR Gold Buffer, 2.5 mM MgCl₂, and 0.5 units AmpliTaq Gold (Perkin-Elmer, Norwalk, CT) DNA polymerase in a 10-μL reaction solution. PCR was carried out in a thermal cycler with the following profile: 7 min at 95°C, followed by 15 s at 94°C for denaturation, 30 s at 55°C for annealing and 2 min at 72°C for synthesis. The PCR amplification was continued for 40 cycles with a final extension for 30 min at 72°C. Samples were stored at -20°C until needed.

Capillary electrophoretic analysis of fluorescently-labeled, amplified DNA markers were visualized as peaks on electropherograms with the automated ABI PRISM 310 Genetic Analyzer (Perkin-Elmer Corp., Norwalk, CT) equipped with genotyping and Genescan analysis software (PE Applied Biosystems, Foster City, CA, USA). Because of the high sensitivity of the system, a DNA marker was considered valid if it had a peak height of at least 100 fluorescent units and plus or minus one base size difference with the nearest DNA fragment peak. Fluorescently labeled PCR products (1:30 diluted in sterile water) were mixed in 10 μL formamide (AMRESCO Inc., Solon, OH) and 0.2 μL of an internal size standard DNA labeled with a ROX dye, denatured at 95°C for 5 min, kept on ice at least for 4 min and loaded on the automated ABI PRISM 310 Genetic Analyzer. Variation in the amplified products was compared with an internal size standard DNA labeled with the ROX dye different from that of the sample. Computer-assisted analysis of the data was performed with

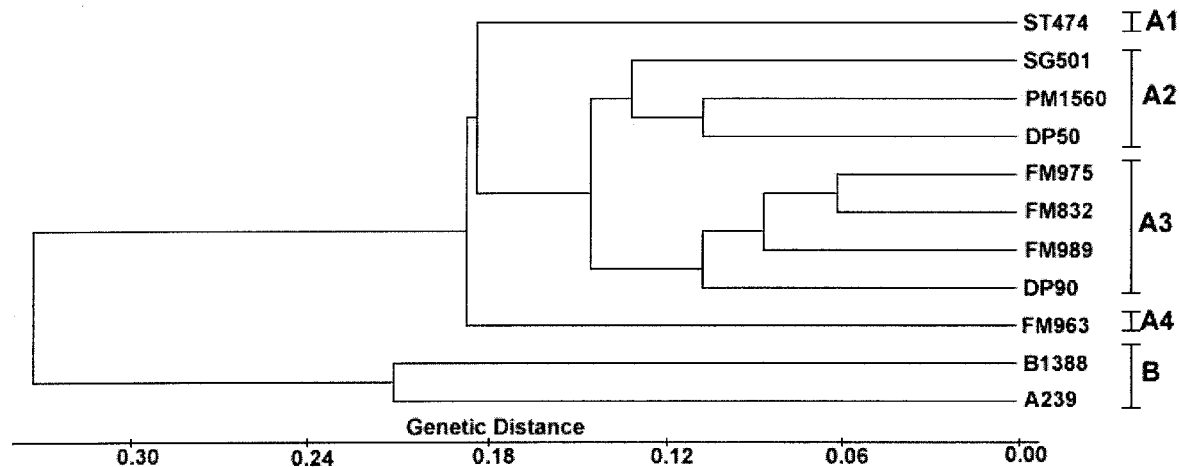


Fig. 1. Dendrogram presenting the association of 11 cotton genotypes determined by Unweighted Paired Group Method using Arithmetic Averages (UPGMA) cluster analysis of 69 polymorphic SSR marker loci.

GeneScan software using local southern method. Additional parameters in the ABI 310 system were injection time 10 to 15 s, electrophoresis voltage 13 kV, injection voltage 15 kV, collection time 26 min, run temperature 60°C, and syringe pump time 180 s.

Data Analysis

Final analyses were performed on a total of 90 primer pairs. Peaks, representative of bands in the electropherograms were then coded as 1 or 0 for presence or absence of the band, respectively. Genetic distances between all pairs of parents were calculated with the PAUP* 4.0b5 software (Swofford, 2000) as $GD = 1 - S_{xy}$ according to the method developed by Nei and Li (1979):

$$S_{xy} = 2 n_{xy} / (n_x + n_y),$$

where S_{xy} = measure of genetic similarity between pairs of parents, n_{xy} = number of bands common in parents X and Y , n_x = number of bands in parents X , and n_y = number of bands in parents Y .

A phylogenetic tree (Fig. 1) and matrix of GD (Table 2) were constructed by the Unweighted Pair Group Mean Average (UPGMA) method of Saitou and Nei (1987) by PAUP* 4.0b5 software (Swofford, 2000). Heterosis of F₂-bulk population for all traits was calculated by comparing each F₂-bulk population with the respective midparent (F₂-MP) mean per replication at each location.

Environments were considered as a combination of years and locations. Combined analysis of variance over environments revealed significant genotype \times environment interactions. Therefore, analyses of variance by individual environments were performed with the MIXED procedure of SAS

(SAS Institute, 2000) for each set of data to test for significant differences ($P < 0.05$) among the parents and F₂-bulk populations. Replications were considered random effects. Entries (Parents and F₂s) were considered fixed effects. Least Significant Differences (LSD) were calculated for pair-wise comparisons among entries (i.e., parents and F₂ bulk) in each environment by means of the restricted maximum likelihood (REML) estimator of the corresponding standard error of the mean differences (Littell et al., 1996). Entry means over replications within environments were used in correlation analyses with GD. Correlations between GD and the agronomic and fiber traits were estimated by PROC CORR of SAS (SAS Institute, 2000). For this analysis, data was divided into six subsets as follows: Set 1—U.S. cultivars \times U.S. cultivars (US \times US), Set 2—Australian cultivars \times Australian cultivars (AU \times AU), Set 3—Australian cultivars \times U.S. cultivars (AU \times US), Set 4—U.S. cultivars \times day-neutral converted lines (US \times day neutral), Set 5—Australian cultivars \times day-neutral converted lines (AU \times day neutral), and Set 6—all crosses.

RESULTS AND DISCUSSION

Genetic Diversity among the Parental Lines

Sixty of the 90 SSR primer pairs amplified 69 polymorphic marker loci among 11 cotton genotypes of diverse genetic background (Table 1). Nine primer pairs (CML66, BNL530, BNL1417, BNL1551, BNL1897, BNL2440, BNL2448, BNL2646, and BNL3599) amplified two loci each and 51 primer pairs (CML43, CML63, CML68, BNL119, BNL169, BNL193, BNL256, BNL285, BNL409, BNL597, BNL632, BNL840, BNL1053,

Table 2. Genetic distance coefficients calculated for 10 cotton genotypes from 69 SSR marker loci.

Genotypes	A239	B1388	DP50	DP90	FM832	FM963	FM975	FM989	PM1560	SG501
B1388	0.21†									
DP50	0.27	0.27								
DP90	0.31	0.26	0.15							
FM832	0.30	0.30	0.12	0.12						
FM963	0.31	0.31	0.18	0.14	0.19					
FM975	0.30	0.28	0.12	0.09	0.06	0.16				
FM989	0.29	0.28	0.10	0.11	0.07	0.17	0.09			
PM1560	0.28	0.26	0.10	0.17	0.17	0.19	0.15	0.15		
SG501	0.33	0.30	0.14	0.16	0.16	0.18	0.13	0.14	0.12	
ST474	0.31	0.34	0.16	0.22	0.18	0.19	0.17	0.16	0.15	0.16

† Genetic Distance coefficients calculated according to the formula $GD = 1 - S_{xy}$.

BNL1059, BNL1064, BNL1162, BNL1231, BNL1414, BNL1423, BNL1440, BNL1495, BNL1672, BNL1681, BNL1694, BNL1707, BNL2499, BNL2530, BNL2590, BNL2597, BNL2895, BNL2986, BNL3065, BNL3084, BNL3090, BNL3264, BNL3279, BNL3280, BNL3368, BNL3383, BNL3400, BNL3441, BNL3442, BNL3479, BNL3482, BNL3649, BNL3792, BNL3816, BNL3895, BNL3955, BNL3971, and BNL3994) amplified one locus each. Similar results were previously obtained by Liu et al. (2000a, b) who mapped many of these SSR loci to different chromosomes. A total of 139 different alleles were amplified by 60 SSR primer pairs yielding an average of two SSR alleles per marker locus. The SSR markers used in this study were reported to be present on at least 13 of the 26 chromosomes of the cotton genome. Approximately 55% of the SSR markers have been assigned to the A genome and 46% to the D genome of tetraploid cotton.

Genetic distance coefficients based on the polymorphic SSR marker loci ranged from 0.06 to 0.34 for the 11 parental lines (Table 2). The highest GD (0.34) was detected between the cultivar ST474 and the day-neutral converted line B1388. The lowest GD (0.06) was observed between Australian cultivars, FM832 and FM975. The GD between cultivars and day-neutral converted lines ranged from 0.26 to 0.34. Among the U. S. cultivars, GD ranged from 0.10 to 0.22 while GD of Australian cultivars varied from 0.06 to 0.19, indicating the presence of a narrow genetic base among the Australian and U.S. cultivars. Multani and Lyon (1995) observed GD of 0.01 to 0.08 among nine Australian cultivars which also showed a lack of genetic diversity. Iqbal et al. (1997) also found very high genetic similarity of 0.82 to 0.93 among 17 *G. hirsutum* cultivars on the basis of random amplified polymorphic DNA (RAPD) markers. On the basis of SSR markers, Ulloa et al. (1999) observed a GD of 0.18 within Acala and Delta cottons while the GD within the Pima PS series was approximately 0.16. The monoculture of a few successful cotton cultivars and the extensive use of them as parents in breeding programs has limited genetic diversity, possibly contributing to yield stagnation in the 1990s (Van Esbroeck et al., 1998). Furthermore, Bowman (2000) suggested that the present situation of high genetic uniformity in cotton is unlikely to change. He reported that genetic uniformity of cotton cultivars reflects reselection within cultivars, disregard for coefficient of parentage in commercial breeding programs, and reduced efforts in germplasm enhancement. Low GD values indicate that it is necessary to introgress new alleles into the U.S. cotton germplasm base to enhance its genetic diversity.

The GD between day-neutral converted lines of wild race stocks, A239 and B1388, was 0.21. This is somewhat surprising because their recurrent parental accessions were collected in different countries, and they have different donor parents for the day-neutral flowering trait (DP16 and DES56, respectively). These results are in agreement with the findings of Liu et al. (2000a), who reported a very narrow genetic base in a subset of day-neutral selections from crosses of race-stock accessions and upland cultivars compared with the *G. hirsutum*

genetic standard TM-1. The lack of diversity between A239 and B1388 could be due to the fact that these accessions were selected for the same traits of insensitivity to photoperiod and high fiber strength. During this process, large chromosomal blocks surrounding genes for fiber strength, earliness, and day neutrality may be introgressed as common alleles in both lines. Lack of enough recombination in portions of the cotton genome could be also responsible for the small GD.

The UPGMA tree generated from genetic distance coefficients, grouped the 11 cotton genotypes into two major clusters (Fig. 1). Cluster A (denoted as Subclusters A1-A4 in Fig. 1) represents the cultivars, and Cluster B the day-neutral converted lines. Within Cluster A, there were four subclusters, A1 (ST474), A2 ('SG501', 'PM1560', and 'DP50'), A3 (FM975, FM832, 'FM989', and 'DP90'), and A4 ('FM963') (Fig. 1). Although PM1560, SG501, and ST474 have one parent in common, 'DES119' (Calhoun et al., 1994) (Table 1), ST474 was in a separate subcluster. DP 90 was included with the Australian cultivars in the A3 cluster. DP 90 was more closely related to the Australian cultivars compared with the U.S. cultivars indicating that the Australian cultivars and DP 90 may have a similar genetic background, or DP 90 is a frequent parent in the pedigree of Australian cultivars. Subcluster A4 consists only of the Australian cultivar, FM 963. The Australian cultivars were distinguished by clustering close together compared with the U.S. cultivars, indicating their origin from similar genetic backgrounds. Long-term selection for similar agronomically desirable traits may have caused the cultivars to be genetically more uniform and thus to cluster close together. Even though pedigree information on the Australian cultivars was not available, the majority of the groups resulting from cluster analysis were as expected on the basis of their pedigree and geographic information which further supports the strength of SSR markers for the detection of GD.

Relationship of Genetic Distance with F₂-Bulk Performance

There were significant differences ($P < 0.001$) among entries (i.e., parents and F₂ bulk) for all agronomic and fiber traits in each environment. There were changes in the rank of the crosses from environment to environment. The number of F₂-bulk population means differing from environment to environment demonstrated the strong impact of environment on fiber properties, yield, and yield components (Table 3). The effects of severe dry conditions in Environments 3 and 4 during 2000 perhaps affected the performance of F₂-bulk populations, since several of them performed better than their mid-parents in agronomic and fiber traits.

Correlations between agronomic and fiber traits of F₂-bulk populations and GD across environments varied in magnitude and direction from -0.99 to 0.90 (Table 4). Midparent heterosis for lint percentage and fiber elongation were significantly correlated with GD in all environments for Set 1. Lint percentage and mid-parent heterosis for seed cotton yield and lint yield were signifi-

Table 3. Number of F₂-bulk population means out of 54 significantly higher than the mid-parent (\pm LSD) for yield components and fiber properties grown at four environments in Mississippi during 1999 and 2000.

Traits	Environment 1	Environment 2	Environment 3	Environment 4
Seed cotton yield	10	18	42	42
Lint yield	9	12	38	40
Lint Percentage	6	1	1	1
Boll Size	5	4	6	10
Fiber Micronaire	6	3	11	6
Fiber Elongation	0	0	0	1
Fiber strength	2	0	0	0
Fiber 50% span length	7	5	1	6
Fiber 2.5% span length	12	14	15	10

cantly correlated with GD in Set 3 in most of the environments. Fiber micronaire showed significant correlation with GD in Set 5 in two environments. Finally, in Set 6, lint yield, lint percentage, fiber strength, and 2.5%

fiber span length were significantly correlated with GD in all environments. Seed cotton yield, mid parent-heterosis for seed cotton yield, lint yield and boll size, fiber micronaire, fiber elongation, and 50% fiber span length

Table 4. Correlations between agronomic traits of F₂-bulk populations and GD for six sets of parental cotton genotypes.

Environment	Trait	Set 1†	Set 2	Set 3	Set 4	Set 5	Set 6
1	Seed cotton yield	0.08	0.39	0.23	-0.25	-0.04	-0.57***
2		-0.32	-0.12	0.27	-0.10	0.03	-0.23
3		-0.49	-0.26	0.04	0.58	0.90**	-0.62***
4		-0.46	-0.21	0.09	0.36	-0.59	-0.61***
1	MPH‡ seed cotton yield	0.34	0.79	0.68***	-0.07	0.17	0.36**
2		0.14	0.05	0.60**	0.47	0.23	0.48***
3		-0.18	0.97	0.14	-0.24	0.87*	0.30*
4		0.41	0.40	0.54*	0.53	0.65	0.26
1	Lint yield	0.18	0.34	0.29	-0.14	-0.01	-0.67***
2		-0.14	-0.09	0.39	0.03	0.01	-0.43***
3		-0.30	-0.23	0.17	0.68*	0.89**	-0.71***
4		-0.31	-0.05	0.17	0.61	-0.59	-0.70***
1	MPH lint yield	0.24	0.81*	0.67***	-0.11	0.27	0.28*
2		0.05	0.06	0.61**	0.48	0.21	0.44***
3		-0.30	0.98	0.11	-0.41	0.89**	0.23
4		-0.53	0.44	0.48	0.50	0.70	0.25
1	Lint percentage	0.35	-0.46	0.25	0.32	0.23	-0.77***
2		0.45	0.42	0.48*	0.54	-0.21	-0.75***
3		0.34	0.20	0.50*	0.45	0.07	-0.72***
4		0.39	0.40	0.55**	0.60	-0.34	-0.68***
1	MPH lint percentage	-0.71*	0.67	0.04	-0.29	0.41	0.20
2		-0.69*	0.14	0.11	-0.25	-0.30	0.12
3		-0.78**	-0.99**	0.13	-0.35	-0.37	0.20
4		-0.72*	0.37	-0.22	0.06	-0.64	0.27
1	Boll size	-0.40	-0.56	0.13	-0.08	-0.09	-0.20
2		-0.13	-0.11	0.31	0.21	0.33	-0.18
3		-0.55	-0.49	0.05	0.31	0.35	-0.33**
4		0.54	-0.35	0.17	0.24	0.22	-0.15
1	MPH boll size	0.57	0.13	0.17	-0.11	-0.37	0.22
2		0.58	0.05	0.40	0.06	0.03	0.17
3		0.01	0.03	-0.56*	0.14	-0.42	0.33*
4		0.01	-0.79	0.43	0.30	-0.84*	0.29*
1	Fiber micronaire	-0.43	-0.42	-0.43*	-0.16	-0.61	0.23
2		-0.36	0.12	0.03	-0.02	-0.10	0.36**
3		0.09	0.14	-0.14	0.36	-0.70*	0.55***
4		0.15	0.11	-0.23	0.42	-0.75*	0.54***
1	Fiber elongation	-0.75*	0.32	0.24	-0.53	-0.03	-0.32**
2		-0.68*	-0.21	0.22	-0.61	-0.81**	-0.26
3		-0.85**	0.35	0.27	-0.48	0.43	-0.37**
4		-0.85**	0.49	0.40	-0.11	-0.18	-0.48***
1	Fiber strength	0.27	-0.76	-0.22	0.30	-0.72*	0.46***
2		0.33	-0.61	-0.29	-0.10	-0.46	0.38***
3		0.49	-0.66	-0.37	-0.10	-0.53	0.32**
4		0.12	-0.46	-0.18	0.06	-0.30	0.36**
1	Fiber 50% span length	-0.18	-0.16	-0.09	-0.38	-0.43	-0.45***
2		-0.08	-0.57	0.38	-0.21	-0.51	-0.42***
3		0.01	-0.46	-0.28	-0.13	-0.51	0.13
4		-0.64	-0.90**	0.14	-0.12	-0.45	-0.15
1	Fiber 2.5% span length	-0.13	-0.54	-0.18	-0.40	-0.40	-0.64***
2		0.06	-0.76	-0.07	-0.16	-0.54	-0.68***
3		-0.14	-0.92*	-0.35	-0.05	-0.50	-0.46***
4		-0.22	-0.74	-0.30	-0.30	-0.42	-0.54***

* Indicates significance at $P = 0.05$.

** Indicates significance at $P = 0.01$.

*** Indicates significance at $P = 0.001$.

† Set 1, US \times US; Set 2, AU \times AU; Set 3, AU \times US; Set 4, US \times day-neutral; Set 5, AU \times day-neutral; Set 6, all 54 crosses.

‡ Mid-parent heterosis (MPH).

were also significantly correlated with GD in most of the environments in Set 6.

Correlations between agronomic and fiber traits of F_2 -bulk populations and GD were not consistent across environments and sets. Therefore, they were not generally useful for F_2 performance prediction. However, it was a better predictor of traits such as mid parent-heterosis for seed cotton yield and fiber strength in Set 6 whose genotypes exhibited a maximum range of GD values. Our findings in cotton agree with Melchinger (1993) who concluded that GD estimates based on markers randomly arranged across the maize genome were of no value in predicting hybrid performance. Additionally, Bernardo (1992) stated that molecular marker heterozygosity would be most valuable for predicting hybrid performance in crop species under conditions such as strong dominance effects, and high trait heritability, while, Charcosset et al. (1991) anticipated the need for linkage disequilibrium between marker loci and quantitative trait loci (QTL) for marker heterozygosity and heterosis to be associated. In cotton, some of these conditions may not be present, thus affecting the ability of markers to predict F_2 -bulk population performance. For instance, we did not apparently have sufficient linkage between markers and QTLs to predict population performance. Furthermore, correlations of GD based on marker heterozygosity with performance and heterosis differed from one trait to another and depended on the genetic background of the germplasm (Lee et al., 1989; Godshalk et al., 1990; Melchinger et al., 1990; Boppenmaier et al., 1993; Zhang et al., 1996; Xiao et al., 1996; Saghai Maroof et al., 1997). In this study, the direction of the correlations often changed among crosses and environments. In Set 6, seed cotton yield, lint yield, lint percentage, fiber strength, and 2.5% fiber span length were the only traits for which the direction of the correlation between GD and F_2 population performance was the same in all environments. Midparent heterosis for seed cotton yield was positively correlated with GD in Set 3 and 6; however, lint percentage was negatively correlated with GD in Set 6 and positively correlated in Set 3. A negative association between GD and mid parent heterosis for lint percentage was observed in Set 1 in all environments. Perhaps heterosis was expressed in some traits at some environments for certain genetic backgrounds.

In regards to fiber properties, significant negative correlations for micronaire reading were observed only in Set 5 in two environments indicating that day-neutral converted line A239, which has a low micronaire reading, could be used to develop lower micronaire reading germplasm. This also agrees with Cheatham (2001) who observed that predicted micronaire values for crosses with A239 decreased in advanced generations of selection. A negative correlation was observed between fiber elongation and GD in Set 1 despite parents with desirable fiber elongation.

In summary, commercial cultivars exhibited little genetic diversity with GD ranging from 0.06 to 0.22. The GD between the day-neutral converted lines A239 and B1388 was 0.21, which was similar to the greatest GD

between commercial cultivars. The day-neutral converted lines, A239 and B1388, had lower yield, lower lint percentage, and shorter and stronger fibers than the cultivars, but their fiber strength were nearly as high as that of FM 832. Crosses of A239 and B1388 with the cultivars should gain alleles for higher fiber strength, but to use the strength genes from A239 and B1388, careful attention should be given to fiber length, lint percentage, and lint yield to avoid negatively affecting these traits.

An assessment of the usefulness of SSR markers in breeding cotton for yield improvement and fiber quality needs further consideration. We must acknowledge that a limited number of SSR markers were used to make this experiment cost effective. Perhaps more SSR markers covering all 26 chromosomes and at higher density are needed to better predict F_2 population performance with GD. Improving the association between SSR marker diversity and F_2 -bulk population performance also would require SSR markers linked to QTLs for agronomic traits and fiber properties. Our results also demonstrated that the performance of F_2 -bulk populations does not always depend on the GD, but on the genetic background of the parental germplasm. Genetic diversity has to be incorporated in any breeding program with a proper plan, so that maximum potential can be achieved by incorporating favorable QTLs in a crossing program and removing the unfavorable ones from parental lines (Liu and Wu, 1998). Finally, our study also revealed very low genetic diversity among improved commercial cultivars and, thus the need for new germplasm introgression.

ACKNOWLEDGMENTS

The authors express thanks to Mr. Douglas Dollar, Dr. Mehmet Karaca, Ms. Chethana Bezawada, and Mr. Russ Hayes for their technical assistance. We also thank Drs. Roy G. Cantrell, Michael Lee, James B. Holland, Mauricio Ulloa, Lloyd May, and three anonymous reviewers for their critical evaluation of the manuscript and helpful comments. Finally, we are grateful to Dr. Clarence E. Watson, Jr., for his valuable assistance with the statistical analyses.

REFERENCES

- Bernardo, R. 1992. Relationship between single-cross performance and molecular marker heterozygosity. *Theor. Appl. Genet.* 83:628–634.
- Boppenmaier, J.A., E. Melchinger, G. Seitz, H.H. Geiger, and R.G. Herrmann. 1993. Genetic diversity for RFLPs in European maize inbreds. III. Performance of crosses within versus between heterotic groups for grain traits. *Plant Breed.* 111:217–226.
- Bowman, D.T. 2000. Attributes of public and private cotton breeding programs. *J. Cotton Sci.* 4:130–136.
- Bowman, D.T., O.L. May, and D.S. Calhoun. 1996. Genetic base of upland cotton cultivars released between 1970 and 1990. *Crop Sci.* 36:577–581.
- Calhoun, D.S., D. T. Bowman, and O. L. May. 1994. Pedigrees of upland and pima cotton cultivars released between 1970 and 1990. *Bull.* 1017. *Miss. Agric. and For. Exp. Stn.* Mississippi State, MS.
- Charcosset, A., M. Lefort-Buson, and A. Gallais. 1991. Relationship between heterosis and heterozygosity at marker loci: a theoretical computation. *Theor. Appl. Genet.* 81:571–575.
- Cheatham, C.L. 2001. Genetics and combining ability of yield and fiber properties associated with selected American and Australian

- cotton genotypes. Ph.D. diss. (Diss. Abstr. Int. 02B:608), Mississippi State Univ., Mississippi State, MS.
- Cheres, M.T., J.F. Miller, J.M. Crane, and S.J. Knapp. 2000. Genetic distance as a predictor of heterosis and hybrid performance within and between heterotic groups in sunflower. *Theor. Appl. Genet.* 100:889–894.
- Dudley, J.W., M.A. Saghai Maroof, and G.K. Rufener. 1991. Molecular markers and grouping of parents in maize breeding programs. *Crop Sci.* 31:718–723.
- Godshalk, E.B., M. Lee, and K.R. Lamkey. 1990. Relationship of restriction fragment length polymorphisms to single-cross hybrid performance of maize. *Theor. Appl. Genet.* 80:273–280.
- Iqbal, M.J., N. Aziz, N.A. Saeed, Y. Zafar, and K.A. Malik. 1997. Genetic diversity evaluation of some elite cotton varieties by RAPD analysis. *Theor. Appl. Genet.* 94:139–144.
- Lanza, L.L.B., C.L. de Souza Jr., L.M.M. Ottoboni, M.L.C. Vieira, and A.P. de Souza. 1997. Genetic distance of inbred lines and prediction of maize single-cross performance using RAPD markers. *Theor. Appl. Genet.* 94:1023–1030.
- Lee, M., E.B. Godshalk, K.R. Lamkey, and W.W. Woodman. 1989. Association of restriction fragment length polymorphism among maize inbreds with agronomic performance of their crosses. *Crop Sci.* 29: 1067–1071.
- Li, C.D., B.G. Rosnagel, and G.J. Scoles. 2000. The development of oat microsatellite markers and their use in identifying relationships among *Avena* species and oat cultivars. *Theor. Appl. Genet.* 101:1259–1268.
- Littell, R.C., G.A. Milliken, W.W. Stroup, and R.D. Wolfinger. 1996. SAS system for mixed models. SAS Institute Inc., Cary, NC.
- Liu, S., R.G. Cantrell, J.C. McCarty, Jr., and J.M. Stewart. 2000a. Simple sequence repeat-based assessment of genetic diversity in cotton race stock accessions. *Crop Sci.* 40:1459–1469.
- Liu, S., S. Saha, D. Stelly, B. Burr, and R.G. Cantrell. 2000b. Chromosomal assignment of microsatellite in cotton. *J. Hered.* 91:326–332.
- Liu, X.C., and J.L. Wu. 1998. SSR heterogenic patterns of parents for making and predicting heterosis in rice breeding. *Mol. Breed.* 4:263–268.
- Manjarrez-Sandoval, P., T.E. Carter, Jr., D.M. Webb, and J.W. Burton. 1997. RFLP genetic similarity estimates and coefficient of parentage as genetic variance predictors for soybean yield. *Crop Sci.* 37:698–703.
- Martin, J.M., L.E. Talbert, S.P. Lanning, and N.K. Blake. 1995. Hybrid performance in wheat as related to parental diversity. *Crop Sci.* 35:104–108.
- Melchinger, A.E. 1993. Use of RFLP markers for analysis of genetic relationships among breeding materials and prediction of hybrid performance. p. 621–628. *In* D.R. Buxton et al. (ed.) *International crop science I*. ASSA, CSSA, and SSSA, Madison WI.
- Melchinger, A.E., M. Lee, K.R. Lamkey, and W.L. Woodman. 1990. Genetic diversity for restriction fragment length polymorphisms: relation to estimated genetic effects in maize inbreds. *Crop Sci.* 30:1033–1040.
- Meredith, W.R., Jr. 1990. Yield and fiber-quality potential for second-generation cotton hybrids. *Crop Sci.* 30:1045–1048.
- Meredith, W.R., Jr., and J.S. Brown. 1998. Heterosis and combining ability of cotton originating from different regions of the United States. *J. Cotton Sci.* 2:77–84.
- Multani, D.S., and B.R. Lyon. 1995. Genetic fingerprinting of Australian cotton cultivars with RAPD markers. *Genome* 38:1005–1008.
- Nei, M., and W. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. (USA)* 76:5269–5273.
- Reddy, O. U. K., A.E. Pepper, I. Abdurakhmonov, S. Saha, J.N. Jenkins, T. Brooks, Y. Bolek, and K.M. El-Zik. 2001. New dinucleotide and trinucleotide microsatellite marker resources for cotton genome research. *J. Cotton Sci.* 5:103–113.
- Saghai Maroof, M.A., G.P. Yang, Q. Zhang, and K.A. Gravois. 1997. Correlation between molecular marker distance and hybrid performance in U.S. southern long grain rice. *Crop Sci.* 37:145–150.
- Saha, S., F. E. Callahan, D. A. Dollar, and J. B. Creech. 1997. Effect of lyophilization of cotton tissue on the quality of extractable DNA, RNA and protein. *J. Cotton Sci.* 1:11–14.
- Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406–423.
- SAS Institute. 2000. The SAS system for windows Version 8 (Tsm1). SAS Institute, Cary, NC.
- Senior, L.M., J.P. Murphy, M.M. Goodman, and C.W. Stuber. 1998. Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. *Crop Sci.* 38:1088–1098.
- Smith, J.S.C., E.C.L. Chin, H. Shu, O.S. Smith, S.J. Wall, M.L. Senior, S.E. Mitchell, S. Kresovich, and J. Ziegler. 1997. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): Comparisons with data from RLFPs and pedigree. *Theor. Appl. Genet.* 95:163–173.
- Smith, J.S.C., and O.S. Smith. 1992. Fingerprinting crop varieties. *Adv. Agron.* 47:85140.
- Smith, O.S., J.S.C. Smith, S.L. Bowen, R.A. Tenborg, and S.J. Wall. 1990. Similarities among a group of elite maize inbreds as measured by pedigree, F₁ grain yield, grain yield, heterosis, and RFLPs. *Theor. Appl. Genet.* 80:833–840.
- Swofford, D.L. 2000. PAUP* 4.0 Beta Version. Phylogenetic analysis using Parsimony (* and other methods) Version 4. Sinauer Associates, Sunderland, MA.
- Tang, B., J.N. Jenkins, and J.C. McCarty, Jr. 1992. Genotypic stability of cotton varieties, resistant germplasm, and their F₂ hybrids. p. 583–587. *In* Proc. Beltwide Cotton Prod. Res. Conf. National Cotton Council, Memphis, TN.
- Tang, B., J.N. Jenkins, J.C. McCarty, and C.E. Watson. 1993a. F₂ hybrids of host plant germplasm and cotton cultivars: I. Heterosis and combining ability for lint yield and yield components. *Crop Sci.* 33:700–705.
- Tang, B., J.N. Jenkins, J.C. McCarty, and C.E. Watson. 1993b. F₂ hybrids of host plant germplasm and cotton cultivars: II. Heterosis and combining ability for fiber properties. *Crop Sci.* 33:706–710.
- Tatineni, V., R.G. Cantrell, and D.D. Davis. 1996. Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. *Crop Sci.* 36:186–192.
- Ulloa, M., W.R. Meredith, Jr., R. Percy, and H. Moser. 1999. Genetic variability within improved germplasm of *Gossypium hirsutum* and *G. barbadense* cottons. p. 73. *In* Agronomy abstracts. ASA, Madison, WI.
- Van Esbroeck, G., D.T. Bowman, D.S. Calhoun, and O.L. May. 1998. Changes in the genetic diversity of cotton in the USA from 1970 to 1995. *Crop Sci.* 38:33–37.
- Wendel, J.F., C.L. Brubaker, and A.E. Percival. 1992. Genetic diversity in *Gossypium hirsutum* and the origin of upland cotton. *Am. J. Bot.* 79:1291–1310.
- Xiao, J., J. Li, L. Yuan, S.R. McCouch, and S.D. Tanksley. 1996. Genetic diversity and its relationship to hybrid performance and heterosis in rice as revealed by PCR-based markers. *Theor. Appl. Genet.* 92:637–643.
- Young, N.D. 1999. A cautiously optimistic version for marker-assisted breeding. *Mol. Breed.* 5:505–510.
- Zhang, Q., Z.Q. Zhou, G.P. Yang, C.G. Xu, K.D. Liu, and M.A. Saghai Maroof. 1996. Molecular marker heterozygosity and hybrid performance in indica and japonica rice. *Theor. Appl. Genet.* 93: 1218–1224.